Optimization notes

Ying Jia Lin

National Cheng Kung University

May 3rd, 2021

Directional derivative

From a starting point $\underline{x}_0$ and a given dirction $\underline{u}$:

- $\underline{x}(\lambda) = \underline{x}_0 + \lambda \underline{u}$
  - λ is a scalar.
- $d\underline{x} = \underline{u} d\lambda$
  - For a small change in λ.
- $F(\lambda) = f(\underline{x}_0 + \lambda \underline{u})$

$$\begin{aligned} dF = df &= (\nabla f(\underline{x}))^\top d\underline{x} \\ &= (\nabla f(\underline{x}))^\top \underline{u} d\lambda = \nabla^\top f \underline{u} \lambda \end{aligned}$$

- $\frac{df}{d\lambda} = \nabla^\top f \underline{u}$
  - If f is minimized at $\underline{x}^* = \underline{x}_0 + \lambda \underline{u}$, then:
    - $\nabla f(\underline{x}^*))^\top f \underline{u} = 0$
    - gradient f evaluated at the minimum point is orthogonalto $\underline{u}$.

Weierstrass Theorem

If $f(\underline{x})$ is continuous on a nonempty feasible set that is cloased and bounded, then $f(\underline{x})$ has a global minimum in this set.

- A set S is bounded if for any point $\underline{x}$ in S, we have $\underline{x}^\top \underline{x} < c$
  - c is a finite positive number.

Single-variable unconstrained optimization

- Necessary condition
  - If a function $f(x)$ has a local minimum at $x = x^*$, and $f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$.
- Sufficient condition
  - Suppose $f'(x^*) = f''(x^*) = \cdots = f^{(m-1)}(x^*) = 0$, but $f^{(m-1)}(x^*) \neq 0$, then $f(x^*)$ is:
    - 1. a local minimum if $f^{(m-1)}(x^*) > 0$ and m is even.
    - 2. a local maximum if $f^{(m-1)}(x^*) < 0$ and m is even.
    - 3. neither a maximum nor a minimum if m is odd.

Multi-variable unconstrained optimization (1)

Definition of r^{th} differential of function f:

$$d^r f(\underline{x}^*) = \sum_{i=1}^{n} \sum_{j=1}^{n} \cdots \sum_{k=1}^{n} h_i h_j \dots h_k \frac{\partial^r f(\underline{x}^*)}{\partial x_i \partial x_j \dots \partial x_k}$$

Example
When (order) $r = 2$ and (number of variables) $n = 3$, we have:

$$\begin{aligned} d^2 f(\underline{x}^*) = d^2 f(x_1^*, x_2^*, x_3^*) &= \sum_{i=1}^{3} \sum_{j=1}^{3} h_i h_j \frac{\partial^2 f(\underline{x}^*)}{\partial x_i \partial x_j} \\ &= h_1^2 \frac{\partial^2 f(\underline{x}^*)}{\partial x_1^2} + h_2^2 \frac{\partial^2 f(\underline{x}^*)}{\partial x_2^2} + h_3^2 \frac{\partial^2 f(\underline{x}^*)}{\partial x_3^2} \\ &+ 2h_1 h_2 \frac{\partial^2 f(\underline{x}^*)}{\partial x_1 \partial x_2} + 2h_2 h_3 \frac{\partial^2 f(\underline{x}^*)}{\partial x_2 \partial x_3} + 2h_1 h_3 \frac{\partial^2 f(\underline{x}^*)}{\partial x_1 \partial x_3} \end{aligned}$$

Multi-variable unconstrained optimization (2)